**Original Research**

# Genetic Evidence for Causal Relationships Between Circulating Cathepsin Levels and Chronic Obstructive Pulmonary Disease: A Mendelian Randomization Study

Chao Duan[1] Ao Zhang[2] Suyan Tian[3]

[1]Intensive Care Unit, The First Hospital of Jilin University, Changchun, Jilin, China

[2]Department of Neurology, The First Hospital of Jilin University, Changchun, Jilin, China

[3]Division of Clinical Research, The First Hospital of Jilin University, Changchun, Jilin, China

***Address correspondence to:***
Suyan Tian
Division of Clinical Research
The First Hospital of Jilin University
Changchun, Jilin, China
Email: wmxt@jlu.edu.cn

***Running Head:*** **Relationship Between Circulating Cathepsin Levels and COPD**

***Keywords*****:** causal inference; cathepsin; chronic obstructive pulmonary disease; Mendelian randomization

***Abbreviations*****:**

***Funding Support***: This study was supported by the Health Department of Jilin Province (2023LC008).

***Date of Acceptance***: July 26, 2025 | ***Published Online***: August 6, 2025

***Citation***: Duan C, Zhang A, Tian S. Genetic evidence for causal relationships between circulating cathepsin levels and chronic obstructive pulmonary disease: a Mendelian randomization study. *Chronic Obstr Pulm Dis*. 2025; Published online August 6, 2025. https://doi.org/10.15326/jcopdf.2025.0626

***This article has an online supplement.***

**Abstract**

**Background:** Cathepsins, a family of lysosomal proteolytic enzymes, have been implicated in the pathogenesis of various complex diseases, including chronic obstructive pulmonary disease (COPD). However, the causal relationship between cathepsins and COPD remains unclear.

**Methods:** This study employed Mendelian randomization (MR) to investigate the potential causal effects of cathepsin levels on COPD risk. This MR analysis utilized genetic data from individuals of European ancestry in the INTERVAL study and FinnGen consortium. Specifically, summary-level genetic data for nine cathepsins (B, E, F, G, H, O, L2, S, and Z) were obtained from the INTERVAL study, while COPD summary statistics were sourced from the FinnGen consortium. We conducted comprehensive MR analyses, including univariable MR, reverse MR, multivariable MR and MR LASSO, to assess causal relationships between cathepsin levels and COPD risk.

**Results:** Univariable MR analysis revealed no significant causal relationships (forward or reverse) between the nine cathepsins and COPD risk. However, multivariable MR analysis identified cathepsins O and S as having direct causal effects on COPD. For cathepsins O and S, OR was estimated as 1.130 ($p = 0.022$, 95% CI = 1.018–1.255) and 1.068 ($p = 0.025$, 95% CI = 1.008–1.132), respectively. Furthermore, these two cathepsins were independent risk factors for COPD after adjusting for smoking.

**Conclusion:** To our knowledge, this is the first MR study to systematically evaluate the causal role of cathepsins in COPD. Further research, particularly clinical trials, is warranted to validate these associations and explore the therapeutic potential of targeting cathepsins in COPD management.




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

## Introduction

Chronic obstructive pulmonary disease (COPD) is a progressive inflammatory lung disorder characterized by persistent airflow limitation [1,2]. The disease clinically manifests through respiratory symptoms including dyspnea, chronic cough, excessive sputum production, and wheezing [2]. Pathologically, COPD encompasses two principal components: emphysema, involving alveolar sac destruction leading to impaired gas exchange, and chronic bronchitis, featuring persistent bronchial inflammation with mucus hypersecretion [3]. The alveolar tissue degeneration in emphysema reduces lung elasticity, while chronic bronchitis causes airway narrowing through inflammatory thickening and mucus plugging. Despite therapeutic advances, COPD remains incurable, with treatment focusing primarily on symptom management and disease progression delay [1]. These clinical realities underscore the critical importance of early detection and sustained intervention, which can substantially improve patient outcomes and quality of life.

Cathepsins are a family of proteolytic enzymes predominantly found in all animals in addition to other organisms. There are several types of cathepsins, each categorized mainly by their enzymatic nature and substrate specificity, including cathepsins B, D, L, S, K, G, H, V, and C among others [4]. These enzymes are known for their roles in lysosomal degradation, where they contribute to breakdown of proteins inside the lysosomes, critical cellular organelles responsible for digesting various biomolecules. Their broad range of functions in cellular maintenance and regulation renders specific cathepsins crucial for cellular functioning and highlights their potential roles in the management and targeted treatments of various complex diseases [5-7].

Growing evidence implicates cathepsins in COPD development through multiple pathological mechanisms [8]. Cathepsins S, L and K contribute to disease progression by degrading extracellular matrix (ECM) components, including elastin and collagen, in lung tissue. This proteolytic activity drives the structural remodeling of airways and parenchyma characteristic of COPD [9–13]. Beyond ECM degradation, cathepsins participate in COPD-associated inflammation by activating cytokines and chemokines that sustain pulmonary inflammatory cascades. Their pathogenic role is further supported by elevated levels detected in COPD patients' sputum and bronchoalveolar lavage fluid [13,14]. Cathepsins also regulate immune responses frequently dysregulated in COPD. For example, cathepsin G modulates neutrophil function [10] – a critical defense mechanism in lungs. Abnormal cathepsin G activity may impair neutrophil responses, exacerbating inflammation and tissue damage [8]. Furthrmore, increased cathepsin E protein in the lung epithelium of COPD patients have been observed [15]

While these observations derive primarily from in vitro and observational studies (which are susceptible to confounding and reverse causality), they collectively suggest cathepsins as potential therapeutic targets. However, the causal relationship between cathepsin levels and COPD risk requires validation through robust methods.




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

Mendelian randomization (MR) is a research method used in epidemiology to assess the causal relationship between potentially modifiable risk factors and health outcomes [15]. This technique leverages genetic variation as a surrogate to examine the causal effect of a modifiable exposure on disease in observational data. In other words, it uses genetic variants as instrumental variables (IVs) to estimate the causal effect of an exposure on the outcome [16,17]. MR relies on the natural random assortment of genes at conception, which obeys Mendel's laws of inheritance. This allocation mimics the randomization process in a randomized controlled trial (RCT), minimizing confounding factors that typically affect observational studies. While MR studies cannot replace RCTs, they offer insights that help bridge the gap between correlation and causation. In this study, we aimed to elucidate the causal relationship between cathepsin levels and COPD risk using MR analyses, and therefore to provide valuable insights on the prevention and early intervention for COPD.

## Methods and Materials

### Experimental data

Genetic association statistics for 9 cathepsin (i.e., cathepsin B, E, F, G, H, O, L2, S, and Z) levels were derived from the INTERVAL study, comprising 3,301 participants of European ancestry [18]. COPD summary statistics were obtained from the FinnGen consortium (https://www.finngen.fi/en), including 6,915 COPD cases and 186,723 controls. Lastly, we used a summary GWAS data for smoking from MRC-IEU consortium, which included 280,508 cases and 180,558 controls.

Instrumental variables for cathepsin levels were selected based on the following criteria: a linkage disequilibrium (LD) threshold of $R^2 < 0.001$ within a 10,000 kb clumping window, and a genome-wide significance level of $5 \times 10^{-6}$.

### Mendelian randomization

For a single nucleotide polymorphism (SNP) to serve as a valid instrumental variable (IV) in Mendelian randomization analyses, three fundamental assumptions must be satisfied: Relevance assumption dictates that the SNP must be strongly correlated with the exposure; Independence assumption requires that the SNP must be independent of any confounders that affect the relationship between the exposure and the outcome; Exclusion restriction assumption requires that the SNP should not have a direct association with the outcome, nor should it be related to the outcome through any pathways other than the exposure.

The inverse-variance weighted (IVW) method served as our primary analytical approach for estimating causal effects, offering optimal statistical power when all genetic variants meet instrumental variable assumptions [19]. However, recognizing that violations of these assumptions—particularly through horizontal pleiotropy—could bias IVW estimates, we implemented supplementary pleiotropy-robust methods to validate our findings. These included


Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

MR-Egger regression, which accounts for balanced pleiotropy [20]; the weighted median approach, providing consistent estimates when even up to 50% of weights derive from invalid instruments [21]; the weighted mode method, effective when the largest SNP cluster shares a common causal estimate [22]; and MR-PRESSO, which identifies and corrects for outlier variants [23].

Both the MR-PRESSO global test and MR-Egger intercept (MR-Egger intercept p-value < 0.05) were used to identify outliers and detect horizontal pleiotropy. Additionally, the MR-PRESSO outlier test was conducted to mitigate or eliminate horizontal pleiotropy by removing outliers (p-value of the MR-PRESSO global test < 0.05). Cochran's Q test was used to assess heterogeneity among SNPs, with a p-value of < 0.05 indicating heterogeneity. If the p-value was less than 0.05, a random-effect model was used to estimate causal effect size. Otherwise, a fixed-effect model was applied instead. The R TwoSampleMR package [24] was utilized for conducting two-sample MR analyses. The MR-PRESSO tests were carried out using the MR-PRESSO package.

Reverse MR analyses in which COPD was considered as the exposure and cathepsins as the outcome were performed to explore reverse causality, using the GWAS studies from the forward MR analyses. Next, multivariable IVW MR analysis involving nine cathepsins as predictors was conducted using the R MendelianRandomization package. To address multi-colinearity, LASSO analysis was performed to select relevant features and construct the final model, and the R MrLasso package (https://github.com/smaityumich/MrLasso) was employed for this analysis. Conditional F-statistics were calculated using the R MVMR package.

Lastly, post-hoc statistical power assessment of the MR analysis was conducted using the mRnd online tool [25] (https://cnsgenomics.com/shiny/mRnd). The current study was designed following the Strengthening the Reporting of Observational Studies in Epidemiology-MR (STROBE-MR) checklist [26].

## Results

### Univariable MR

The causal relationship between nine cathepsins (cathepsin B, E, F, G, H, O, L2, S, and Z) and the risk of COPD was investigated using a two-sample univariable MR analysis. Initially, all p-values from Cochran's Q statistics exceeded 0.05. In conjunction with the results from leave-one-out (LOO) plots (**Supplementary Figure 1**), it was concluded that significant heterogeneity was absent. Furthermore, the MR-Egger intercept test revealed no horizontal pleiotropy, with a p-value greater than 0.05. All F-statistics of single SNPs were larger than 10.

The results of univariable MR analysis are tabulated in **Table 1,** while forest plot in **Figure 1** graphically elucidated IVW results. For example, our analysis showed that the risk of developing COPD for a one-unit increment in abundance level of cethepsin O was estimated to be 1.078 (odds ratio [OR]=1.078, p = 0.081, 95% confidence interval [CI] = 0.991–1.174) by the IVW method. All four complementary MR methods supported in concordant with this null relationship. The OR


Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

was estimated as 1.076 (p=0.478, 95% CI=0.885–1.309) by MR-Egger, 1.033 (p=0.596, 95% CI=0.917–1.163) by weighted median, 1.011 (p=0.900, 95% CI=0.861–1.187) by weighted mode, and 1.044 (p=0.401, 95% CI=0.948–1.149) by MR-PRESSO. Regarding cathepsin S, which has been reported to associate with the development and progression of COPD by numerious prior studies, our analysis found a non-significant postive causal effect of this protein on COPD risk. Specifically, OR was estimated as 1.037 (p=0.464, 95% CI=0.940–1.145) by IVW, 1.064 (p=0.556, 95% CI=0.869–1.302) by MR-Egger, 1.132 (p=0.053, 95% CI=0.999–1.282) by weighted median, 1.140 (p=0.150, 95% CI=0.959–1.356) by weighted mode, and 1.037 (p=0.471, 95% CI=0.989–1.145) by MR-PRESSO, respectively. Overall, none of these nine cathepsins were found to cause an increase or decrease in developing COPD. Furthermore, four MR methods comparison plots are shown in **Supplmentary Figure 2**, demonstrating robust MR results.

Subsequently, reverse MR analysis was performed. Null reverse causal relationship was identified between nine cathepsin types and the risk of developing COPD, which were supported consistently by all MR methods based on corresponding adjusted p-values. For these MR analyses, neither heterogeneity (Cochran's Q *p*-value > 0.05) nor horizontal pleiotropy (MR-Egger intercept *p*-value > 0.05) was detected. The reverse MR results are shown in **Supplementary Table 1.**

**Multivariable MR**

Multivariable MR (MVMR) analyses (including IVW, MR-Egger and MR-PRESSO methods) were performed to analyze the genetic predisposition for nine cathepsin types in relation to the risk of having COPD, which identified both cathepins O and S as a risk factor for COPD (**Table 2**). Specifically, OR was estimated as 1.130 (p=0.022, 95% CI=1.018–1.255) for cathepsin O and 1.068 (p=0.025, 95% CI=1.008–1.132) for cathepsin S by IVW, respectively. Forest plot in **Figure 2** graphically elucidated MVMR IVW analysis results. To address potential multicolinearity between cathepin types and weak instrument bias (some cathespin types have a small conditional F-statistic), we perfomed MR LASSO analysis to select more highly related subtypes, which identified cathepins B, O, S and Z out of nine cathepins. Based on these four cathepins, we redid MVMR analyses (**Table 3**) and the calculated conditional F-statistics for these four cathepins ()indicated marginal weak instrument bias. The results are in line with MVMR analysis result with all nine cathepsin types as covariates.

Furthermore, given that COPD is a heavily smoking-mediated disease, we also performed a MVMR using smoking, cathepins B, O, S and Z as covariates. The analysis results indicated after adjusting for smoking status, both cathepsins O and S were genetically related to COPD risk. Specifically, OR was estimated as 1.217 (p=0.033, 95% CI =1.096–1.458) for cathepsin O and 1.130 (p=0.003, 95% CI =1.043–1.224) for cathepsin S while OR for smoking was as 7.614 (p<0.001, 95% CI =3.819-15.180) by IVW, respectively (**Table 4**).




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Post-hoc power calculation**

Lastly, a post-hoc power calculation was performed to assess the statistical power of the current MR study. At a significance level of 0.05, a sample size of 193,638 individuals including 69,15 cases (3.57%) exhibited a power of 0.80 and 0.56 to detect a 1.2- (OR=1.2) and 1.15-fold risk (OR=1.15) of developing COPD per one-unit increase in the genetically predicted specific cathepsin level. Here, we assumed that these SNPs accounted for approximately 3% of the variance in the cathepsin levels. The sample size of the current MR study was still inadequate to detect a subtle effect.

**Discussion**

In this study, we employed a comprehensive Mendelian randomization (MR) approach incorporating univariable, reverse, and multivariable analyses to investigate potential causal relationships between cathepsin levels and COPD risk. Our multivariable MR results identified cathepsins O and S as potential risk factors for COPD. Existing evidence suggests cathepsin O contributes to COPD pathology through several interconnected pathways. As a protease capable of degrading extracellular matrix (ECM) components, cathepsin O may promote the tissue remodeling and destruction characteristic of COPD. This is supported by observations of elevated cathepsin O levels in COPD patient lung tissues compared to healthy controls [27]. The enzyme's ECM-degrading activity, particularly targeting structural proteins in lung tissue, likely contributes to emphysema development - a hallmark feature of COPD involving alveolar wall destruction that impairs lung elasticity and gas exchange. Beyond its structural effects, cathepsin O appears to influence COPD progression through inflammatory modulation. The enzyme regulates cytokine and chemokine activity, potentially exacerbating the chronic inflammation that drives COPD pathogenesis. Furthermore, emerging evidence positions cathepsin O as a responder to oxidative stress [28], a key driver of COPD development in individuals exposed to cigarette smoke and environmental pollutants. In this context, cathepsin O may participate in processing damaged proteins and organelles resulting from oxidative stress in COPD patients.

Cathepsin S, another key member of the cathepsin protease family, has been consistently shown to be elevated in both lung tissues and serum of COPD patients compared to healthy individuals [7, 29]. As a potent protease, cathepsin S efficiently degrades critical extracellular matrix (ECM) components including elastin and collagen - structural proteins essential for maintaining normal lung architecture[30]. The excessive proteolytic activity of overexpressed cathepsin S may drive the pathological tissue destruction characteristic of emphysema, a defining feature of COPD that involves alveolar wall breakdown and progressive loss of lung elasticity. Beyond its direct effects on parenchymal destruction, cathepsin S-mediated ECM degradation likely contributes to the pathological airway remodeling observed in COPD. By altering the composition and integrity of airway connective tissue, cathepsin S may promote structural changes that lead to airway narrowing and increased stiffness[9-12,31]. These alterations can significantly worsen airflow limitation and contribute to disease progression. Together with cathepsin O, cathepsin S appears to play a multifaceted role in COPD pathogenesis through interconnected mechanisms involving


Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

ECM degradation, inflammatory modulation, tissue remodeling, and oxidative stress responses [32-34]. Their combined actions may create a self-perpetuating cycle of tissue damage and functional decline that characterizes COPD development and progression.

Several limitations should be considered when interpreting our findings. First, the relatively small sample size of the exposure GWAS required us to adopt a more lenient genetic significance threshold Although this enabled the inclusion of additional SNPs, it may have increased susceptibility to weak instrument bias and horizontal pleiotropy. While our comprehensive sensitivity analyses helped mitigate these concerns, residual pleiotropic effects could still influence the results, as is inherent in all Mendelian randomization studies. Second, our analysis was constrained by the availability of GWAS summary data, which included only nine cathepsins and missed several major cathepsin types, including cathepsins A, C, D, K, L and W. Notably, we were unable to evaluate several cathepsins implicated in COPD pathogenesis by previous research, for example, cathepsins D and C [14,27,35]. This limitation not only restricts the comprehensiveness of our investigation but may also introduce biases in the multivariable MR analysis results, as these omitted cathepsins could potentially confound the observed relationships. Another key limitation is that our findings derive from European populations, potentially limiting their applicability to other ethnic groups. Future research must validate these associations in diverse cohorts to determine their broader relevance. Additionally, while MR LASSO was applied to address multi-collinearity, the potential for residual collinearity remains. Lastly, while MR identifies genetic associations between cathepsins and COPD, it cannot assess tissue-specific PTMs regulating cathepsin activation, smoke-induced epigenetic and post-translational regulation, extracellular vs. intracellular cathepsin activity differences, and redox modifications altering protease function. Therefore, future studies should combine proteomics, redox biochemistry, and single-cell analyses to fully elucidate how PTMs and environmental factors (e.g., smoking) modulate cathepsin-driven COPD pathogenesis.

**Conclusion**

To our knowledge, this study represents the first MR study to systematically investigate the causal relationship between circulating cathepsin levels and COPD risk. Our findings suggest that elevated levels of cathepsins O and S may serve as independent risk factors for COPD development, though these observations require further validation.

These findings may provide a novel therapeutic direction for COPD management through targeted modulation of specific cathepsin pathways. However, further investigation - particularly through RCTs - will be crucial to confirm these causal associations and evaluate the clinical potential of cathepsin-focused interventions for COPD patients.




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

## Ethics approval and consent to participate

Ethical approval for this study was waived by the IRB of our institute (First Hospital of Jilin University) as no original research data were collected.

## Consent for publication

Not applicable.

## Availability of data and materials

The GWAS data of cathepsins, smoking, and COPD were downloaded from the (https://gwas.mrcieu.ac.uk). The corresponding GWAS ID numbers are prob-a-718, prob-a-720, prob-a-721, prob-a-723, prob-a-724, prob-a-726, prob-a-727, prob-a-728, prob-a-729, ukb-b-20261, and finn-b-J10_COPD.

## Competing interests

The author declares that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Authors' contributions

ST conceived and designed the experiment; ST and CD ran the analysis and verified the underlying data; CD and ST wrote the original manuscript. AZ, CD, and ST were involved in data interpretation. All authors have read and approved the final version of the manuscript.




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**References**

1. Christenson SA, Smith BM, Bafadhel M, Putcha N. Chronic obstructive pulmonary disease. Lancet. 2022;399(10342):2227-2242.
2. MacLeod M, Papi A, Contoli M, Beghé B, Celli BR, Wedzicha JA, et al. Chronic obstructive pulmonary disease exacerbation fundamentals: Diagnosis, treatment, prevention and disease impact. Respirology. 2021;26(6):532-551.
3. Janjua S, Carter D, Threapleton CJ, Prigmore S, Disler RT. Telehealth interventions: remote monitoring and consultations for people with chronic obstructive pulmonary disease (COPD). Cochrane Database Syst Rev. 2021;7(7):CD013196.
4.Nakanishi H. Cathepsin regulation on microglial function. Biochim Biophys Acta Proteins Proteom. 2020;1868(9):140465.
5. Soond SM, Kozhevnikova MV, Frolova AS, Savvateeva LV, Plotnikov EY, Townsend PA, et al. Lost or Forgotten: The nuclear cathepsin protein isoforms in cancer. Cancer Lett. 2019;462:43-50.
6. Gao S, Zhu H, Zuo X, Luo H. Cathepsin G and Its Role in Inflammation and Autoimmune Diseases. Arch Rheumatol. 2018;33(4):498-504.
7. Brown R, Small DM, Doherty DF, Holsinger L, Booth R, Williams R, et al. Therapeutic Inhibition of Cathepsin S Reduces Inflammation and Mucus Plugging in Adult *β*ENaC-Tg Mice. Mediators Inflamm. 2021;2021:6682657.
8. Pandey KC, De S, Mishra PK. Role of Proteases in Chronic Obstructive Pulmonary Disease. Front Pharmacol. 2017;8:512.
9. Andrault PM, Schamberger AC, Chazeirat T, Sizaret D, Renault J, Staab-Weijnitz CA, et al. Cigarette smoke induces overexpression of active human cathepsin S in lungs from current smokers with or without COPD. Am J Physiol Lung Cell Mol Physiol. 2019;317(5):L625-L638.
10. Effah CY, Drokow EK, Agboyibor C, Ding L, He S, Liu S, et al. Neutrophil-Dependent Immunity During Pulmonary Infections and Inflammations. Front Immunol. 2021;12:689866.
11. Hahn I, Klaus A, Janze AK, Steinwede K, Ding N, Bohling J, et al. Cathepsin G and neutrophil elastase play critical and nonredundant roles in lung-protective immunity against Streptococcus pneumoniae in mice. Infect Immun. 2011;79(12):4893-901.
12. Golovatch P, Mercer BA, Lemaître V, Wallace A, Foronjy RF, D'Armiento J. Role for cathepsin K in emphysema in smoke-exposed guinea pigs. Exp Lung Res. 2009;35(8):631-45.
13. Xu X, Yu T, Dong L, Glauben R, Wu S, Huang R, et al. Eosinophils promote pulmonary matrix destruction and emphysema via Cathepsin L. Signal Transduct Target Ther. 2023;8(1):390.
14. Bchir S, Boumiza S, Ben Nasr H, Garrouch A, Kallel I, Tabka Z, et al. Impact of cathepsin D activity and C224T polymorphism (rs17571) on chronic obstructive pulmonary disease: correlations with oxidative and inflammatory markers. Clin Exp Med. 2021;21(3):457-465.
15. Zhang X, Shan P, Homer R, Zhang Y, Petrache I, Mannam P, etal. Cathepsin E promotes pulmonary emphysema via mitochondrial fission. Am J Pathol. 2014 Oct;184(10):2730-41.
16. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. Res Synth Methods. 2019;10(4):486-496.


Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

17. Larsson SC, Butterworth AS, Burgess S. Mendelian randomization for cardiovascular diseases: principles and applications. Eur Heart J. 2023;44(47):4913-4924.
18. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558:73–9.
19. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37:658–65.
20.Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44:512–25.
21. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40:304–14.
22. Hartwig, FP, Smith, GD, Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int. J. Epidemiol. 2017;46, 1985–98.
23. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50:693–8.
24. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018;7.
25. Brion MJA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol. 2013;42:1497–501.
26. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ. 2021;375:n2233.
27.Singanayagam A, Glanville N, Cuthbertson L, Bartlett NW, Finney LJ, Turek E, et al. Inhaled corticosteroid suppression of cathelicidin drives dysbiosis and bacterial infection in chronic obstructive pulmonary disease. Sci Transl Med. 2019;11(507):eaav3879.
28. Kos J, Mitrović A, Perišić Nanut M, Pišlar A. Lysosomal peptidases-intriguing roles in cancer progression and neurodegeneration. FEBS Open Bio. 2022;12(4):708-38.
29.Small DM, Brown RR, Doherty DF, Abladey A, Zhou-Suckow Z, Delaney RJ, et al. Targeting of cathepsin S reduces cystic fibrosis-like lung disease. Eur Respir J. 2019;53(3):1801523.
30. Brown R, Nath S, Lora A, Samaha G, Elgamal Z, Kaiser R, et al. Cathepsin S: investigating an old player in lung disease pathogenesis, comorbidities, and potential therapeutics. Respir Res. 2020; 21(1):111.
31.Smyth P, Sasiwachirangkul J, Williams R, Scott CJ. Cathepsin S (CTSS) activity in health and disease - A treasure trove of untapped clinical potential. Mol Aspects Med. 2022:101106
32. Rasmussen DG, Sand JM, Karsdal MA, Genovese F. Development of a Novel Enzyme-Linked Immunosorbent Assay Targeting a Neo-Epitope Generated by Cathepsin-Mediated Turnover of Type III Collagen and Its Application in Chronic Obstructive Pulmonary Disease. PLoS One. 2017;12(1):e0170023.




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

33. Lee GH, Cheng NW, Yu HH, Tsai JN, Liu T, Wen ZH, et al. A novel zebrafish model to emulate lung injury by folate deficiency-induced swim bladder defectiveness and protease/antiprotease expression imbalance. Sci Rep. 2019;9(1):12633.
34. Bigot P, Chesseron S, Saidi A, Sizaret D, Parent C, Petit-Courty A, et al. Cleavage of Occludin by Cigarette Smoke-Elicited Cathepsin S Increases Permeability of Lung Epithelial Cells. Antioxidants (Basel). 2022;12(1):5.
35. Chitsamankhun C, Siritongtaworn N, Fournier BPJ, Sriwattanapong K, Theerapanon T, Samaranayake L, et al. Cathepsin C in health and disease: from structural insights to therapeutic prospects. J Transl Med. ;22(1):777.




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Table 1.** Results of univariable Mendelian randomization analysis between cathepsins and chronic obstructive pulmonary disease.

| Type | MR Method | Odds ratio | p-value | MR-Egger intercept (p) | Heterogeneity test (p) | MR-PRESSO global (p) |
|---|---|---|---|---|---|---|
| Cathepsin B | IVW | 0.986 | 0.663 | 0.707 | 0.844 | 0.841 |
| | MR-Egger | 0.960 | 0.602 | | 0.802 | |
| | Weighted median | 1.025 | 0.597 | | | |
| | Weighted mode | 1.033 | 0.546 | | | |
| | MR PRESSO | 0.986 | 0.596 | | | |
| Cathepsin E | IVW | 1.056 | 0.123 | 0.420 | 0.650 | 0.651 |
| | MR-Egger | 1.014 | 0.817 | | 0.630 | |
| | Weighted median | 1.054 | 0.256 | | | |
| | Weighted mode | 1.105 | 0.187 | | | |
| | MR PRESSO | 1.056 | 0.111 | | | |
| Cathepsin F | IVW | 1.024 | 0.665 | 0.442 | 0.569 | 0.572 |
| | MR-Egger | 1.136 | 0.388 | | 0.536 | |
| | Weighted median | 1.031 | 0.677 | | | |
| | Weighted mode | 1.057 | 0.649 | | | |
| | MR PRESSO | 1.024 | 0.648 | | | |
| cathepsin G | IVW | 0.992 | 0.824 | 0.580 | 0.627 | 0.651 |
| | MR-Egger | 0.956 | 0.558 | | 0.572 | |
| | Weighted median | 0.964 | 0.481 | | | |
| | Weighted mode | 0.964 | 0.525 | | | |
| | MR PRESSO | 0.992 | 0.810 | | | |
| cathepsin H | IVW | 0.975 | 0.568 | 0.723 | 0.064 | 0.085 |
| | MR-Egger | 0.946 | 0.572 | | 0.048 | |
| | Weighted median | 0.978 | 0.652 | | | |
| | Weighted mode | 0.949 | 0.343 | | | |
| | MR PRESSO | 0.975 | 0.576 | | | |
| cathepsin O | IVW | 1.078 | 0.081 | 0.982 | 0.683 | 0.179 |
| | MR-Egger | 1.076 | 0.478 | | 0.596 | |
| | Weighted median | 1.033 | 0.596 | | | |




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

| | | | | | | |
|---|---|---|---|---|---|---|
| | Weighted mode | 1.011 | 0.900 | | | |
| | MR PRESSO | 1.044 | 0.401 | | | |
| cathepsin S | IVW | 1.032 | 0.268 | 0.742 | 0.312 | 0.265 |
| | MR-Egger | 1.046 | 0.377 | | 0.268 | |
| | Weighted median | 1.097 | 0.014 | | | |
| | Weighted mode | 1.079 | 0.077 | | | |
| | MR PRESSO | 1.032 | 0.279 | | | |
| cathepsin L2 | IVW | 0.955 | 0.343 | 0.477 | 0.842 | 0.874 |
| | MR-Egger | 0.871 | 0.326 | | 0.824 | |
| | Weighted median | 0.933 | 0.246 | | | |
| | Weighted mode | 0.936 | 0.421 | | | |
| | MR PRESSO | 0.955 | 0.236 | | | |
| cathepsin Z | IVW | 0.985 | 0.626 | 0.565 | 0.850 | 0.832 |
| | MR-Egger | 1.007 | 0.883 | | 0.818 | |
| | Weighted median | 0.979 | 0.600 | | | |
| | Weighted mode | 0.998 | 0.968 | | | |
| | MR PRESSO | 0.985 | 0.539 | | | |




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Table 2.** Results of multivariable Mendelian randomization analysis between cathepsins and chronic obstructive pulmonary disease.

| Type | MR Method | Odds ratio | p-value | MR-Egger intercept p | Heterogeneity test p | PRESSO global p |
|---|---|---|---|---|---|---|
| Cathepsin B | IVW | 0.997 | 0.926 | 0.436 | 0.495 | 0.705 |
| | MR-Egger | 0.984 | 0.666 | | 0.509 | |
| | MR PRESSO | 0.997 | 0.929 | | | |
| Cathepsin E | IVW | 1.043 | 0.221 | | | |
| | MR-Egger | 1.034 | 0.354 | | | |
| | MR PRESSO | 1.035 | 0.338 | | | |
| Cathepsin F | IVW | 1.066 | 0.200 | | | |
| | MR-Egger | 1.068 | 0.193 | | | |
| | MR PRESSO | 1.070 | 0.180 | | | |
| Cathepsin G | IVW | 1.006 | 0.878 | | | |
| | MR-Egger | 1.011 | 0.777 | | | |
| | MR PRESSO | 1.009 | 0.810 | | | |
| Cathepsin H | IVW | 0.950 | 0.195 | | | |
| | MR-Egger | 0.940 | 0.133 | | | |
| | MR PRESSO | 0.949 | 0.186 | | | |
| Cathepsin O | IVW | 1.131 | 0.022 | | | |
| | MR-Egger | 1.148 | 0.012 | | | |
| | MR PRESSO | 1.139 | 0.018 | | | |
| Cathepsin S | IVW | 1.068 | 0.025 | | | |
| | MR-Egger | 1.069 | 0.023 | | | |
| | MR PRESSO | 1.068 | 0.027 | | | |
| Cathepsin L2 | IVW | 0.957 | 0.366 | | | |
| | MR-Egger | 0.946 | 0.258 | | | |
| | MR PRESSO | 0.953 | 0.318 | | | |
| Cathepsin Z | IVW | 0.995 | 0.881 | | | |
| | MR-Egger | 0.994 | 0.851 | | | |
| | MR PRESSO | 0.995 | 0.864 | | | |




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Table 3.** Results of multivariable Mendelian randomization analysis after LASSO feature selection between cathepsins and chronic obstructive pulmonary disease .

| Type | MR Method | Odds ratio | p-value | MR-Egger intercept p | Heterogeneity test p | PRESSO global p |
|---|---|---|---|---|---|---|
| Cathepsin B | IVW | 1.005 | 0.890 | 0.646 | 0.609 | 0.726 |
| | MR-Egger | 0.993 | 0.880 | | 0.572 | |
| | MR PRESSO | 1.005 | 0.886 | | | |
| Cathepsin O | IVW | 1.107 | 0.036 | | | |
| | MR-Egger | 1.109 | 0.033 | | | |
| | MR PRESSO | 1.107 | 0.035 | | | |
| Cathepsin S | IVW | 1.075 | 0.018 | | | |
| | MR-Egger | 1.073 | 0.020 | | | |
| | MR PRESSO | 1.075 | 0.019 | | | |
| Cathepsin Z | IVW | 0.987 | 0.695 | | | |
| | MR-Egger | 0.986 | 0.658 | | | |
| | MR PRESSO | 0.987 | 0.684 | | | |




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Table 4.** Results of multivariable Mendelian randomization analysis between smoking status, cathepsins B, O, S, Z, and chronic obstructive pulmonary disease .

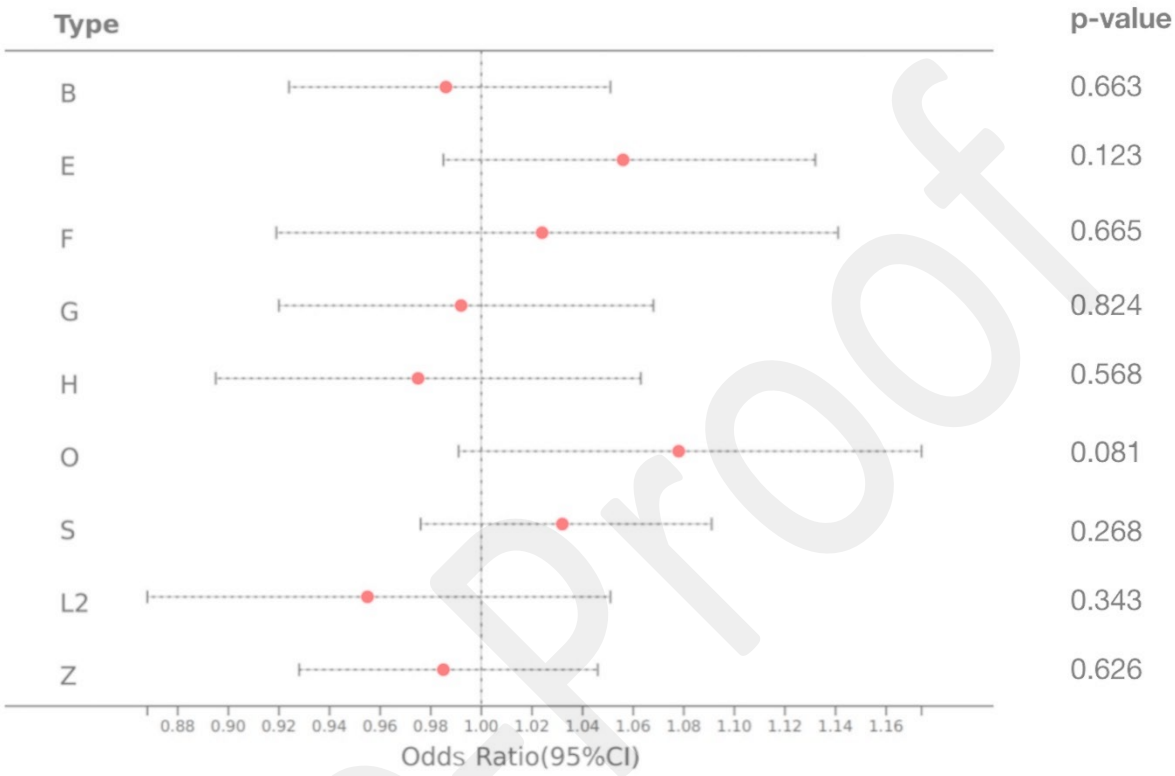
| Type | MR Method | Odds ratio | p-value | MR-Egger intercept p | Heterogeneity test p | PRESSO global p |
|---|---|---|---|---|---|---|
| Cathepsin B | IVW | 1.047 | 0.391 | 0.944 | 0.221 | 0.590 |
| | MR-Egger | 1.049 | 0.444 | | 0.199 | |
| | MR PRESSO | 1.047 | 0.394 | | | |
| Cathepsin O | IVW | 1.217 | 0.033 | | | |
| | MR-Egger | 1.217 | 0.035 | | | |
| | MR PRESSO | 1.217 | 0.036 | | | |
| Cathepsin S | IVW | 1.130 | 0.003 | | | |
| | MR-Egger | 1.129 | 0.003 | | | |
| | MR PRESSO | 1.130 | 0.004 | | | |
| Cathepsin Z | IVW | 1.004 | 0.920 | | | |
| | MR-Egger | 1.004 | 0.920 | | | |
| | MR PRESSO | 1.004 | 0.920 | | | |
| Smoking (ever smoker) | IVW | 7.612 | <0.001 | | | |
| | MR-Egger | 7.594 | <0.001 | | | |
| | MR PRESSO | 7.612 | <0.001 | | | |




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

## Figure Legends

**Figure 1.** Forest plot of univariable Mendelian randomization analysis between the abundance of nine cathepsins (cathepsin B, E, F, G, H, L2, O, S, and Z) and the risk of having chronic obstructive pulmonary disease.

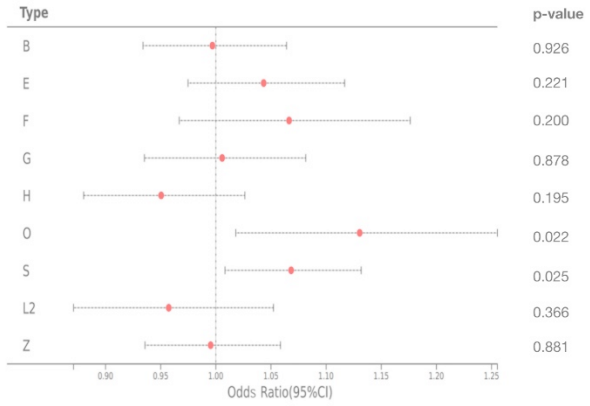






Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Figure 2.** Forest plot of multivariable Mendelian randomization analysis between the abundance of nine cathepsins (cathepsin B, E, F, G, H, L2, O, S, and Z) and the risk of having chronic obstructive pulmonary disease.







Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

## Online Supplement

**Supplementary Table 1.** Results of reverse Mendelian randomization analysis between cathepsins and chronic obstructive pulmonary disease.

| Type | MR Method | Odds ratio | p-value | MR-Egger intercept (p) | Heterogeneity test (p) | MR-PRESSO global (p) |
|---|---|---|---|---|---|---|
| Cathepsin B | IVW | 1.069 | 0.145 | 0.985 | 0.789 | 0.803 |
| | MR-Egger | 1.071 | 0.467 | | 0.744 | |
| | Weighted median | 1.071 | 0.308 | | | |
| | Weighted mode | 1.033 | 0.729 | | | |
| | MR PRESSO | 1.069 | 0.109 | | | |
| Cathepsin E | IVW | 0.933 | 0.133 | 0.786 | 0.679 | 0.681 |
| | MR-Egger | 0.954 | 0.617 | | 0.629 | |
| | Weighted median | 0.964 | 0.579 | | | |
| | Weighted mode | 0.957 | 0.641 | | | |
| | MR PRESSO | 0.933 | 0.116 | | | |
| Cathepsin F | IVW | 1.027 | 0.596 | 0.544 | 0.250 | 0.247 |
| | MR-Egger | 0.973 | 0.786 | | 0.225 | |
| | Weighted median | 1.022 | 0.752 | | | |
| | Weighted mode | 1.012 | 0.909 | | | |
| | MR PRESSO | 1.027 | 0.600 | | | |
| cathepsin G | IVW | 0.930 | 0.112 | 0.670 | 0.585 | 0.600 |
| | MR-Egger | 0.962 | 0.683 | | 0.540 | |
| | Weighted median | 0.952 | 0.453 | | | |
| | Weighted mode | 0.965 | 0.708 | | | |
| | MR PRESSO | 0.930 | 0.109 | | | |
| cathepsin H | IVW | 1.053 | 0.264 | 0.921 | 0.452 | 0.441 |
| | MR-Egger | 1.044 | 0.652 | | 0.398 | |
| | Weighted median | 1.043 | 0.536 | | | |
| | Weighted mode | 1.042 | 0.665 | | | |
| | MR PRESSO | 1.053 | 0.274 | | | |
| cathepsin O | IVW | 1.055 | 0.339 | 0.201 | 0.058 | 0.055 |



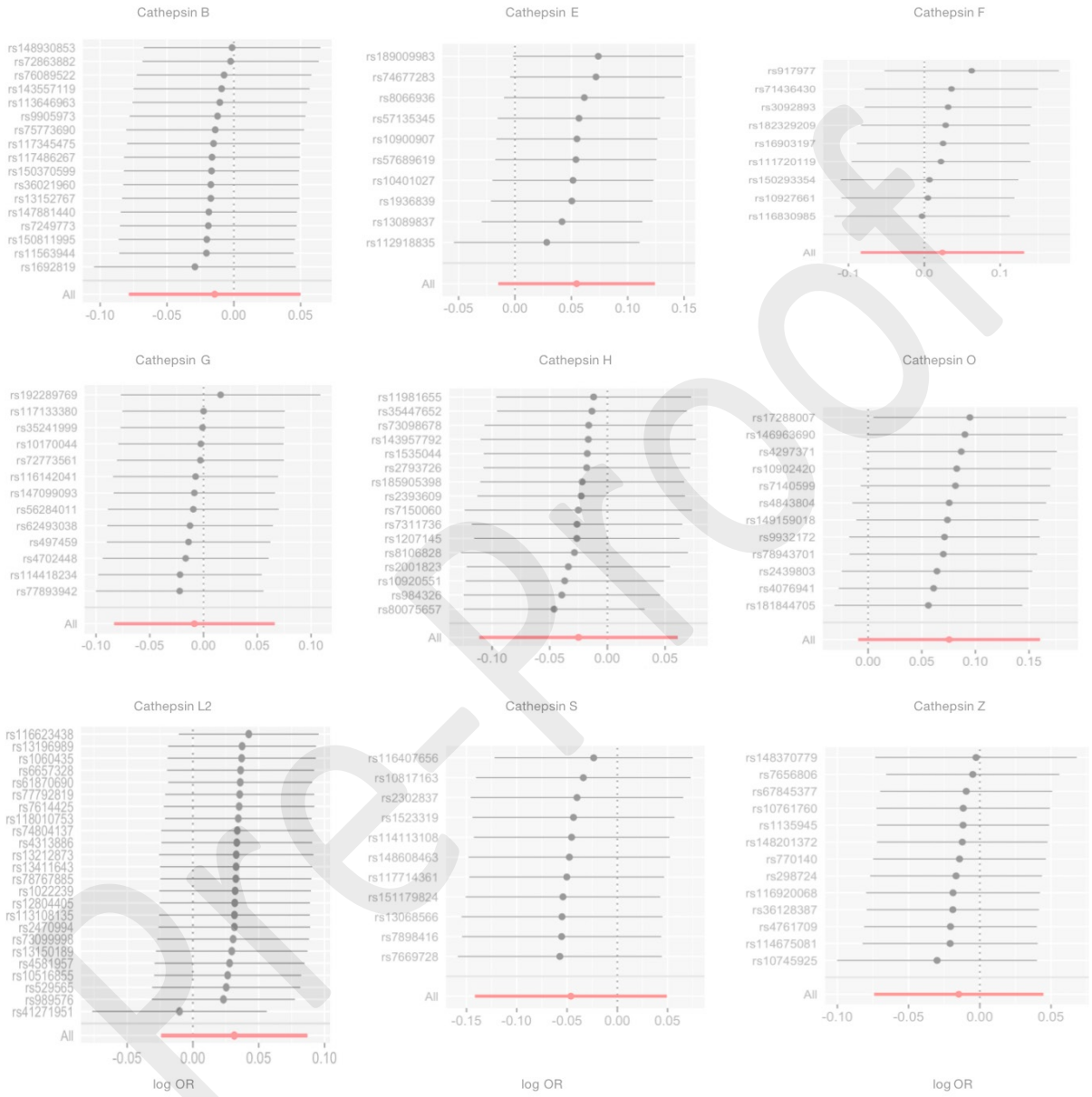

Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

| | | | | | | |
|---|---|---|---|---|---|---|
| | MR-Egger | 1.197 | 0.117 | | 0.076 | |
| | Weighted median | 1.097 | 0.188 | | | |
| | Weighted mode | 1.106 | 0.292 | | | |
| | MR PRESSO | 1.055 | 0.348 | | | |
| cathepsin S | IVW | 1.037 | 0.464 | 0.783 | 0.227 | 0.265 |
| | MR-Egger | 1.064 | 0.556 | | 0.191 | |
| | Weighted median | 1.132 | 0.053 | | | |
| | Weighted mode | 1.140 | 0.150 | | | |
| | MR PRESSO | 1.037 | 0.471 | | | |
| cathepsin L2 | IVW | 1.091 | 0.060 | 0.271 | 0.454 | 0.468 |
| | MR-Egger | 0.996 | 0.966 | | 0.468 | |
| | Weighted median | 1.117 | 0.112 | | | |
| | Weighted mode | 1.097 | 0.356 | | | |
| | MR PRESSO | 1.091 | 0.071 | | | |
| cathepsin Z | IVW | 1.064 | 0.177 | 0.368 | 0.672 | 0.668 |
| | MR-Egger | 1.145 | 0.155 | | 0.667 | |
| | Weighted median | 1.132 | 0.066 | | | |
| | Weighted mode | 1.168 | 0.110 | | | |
| | MR PRESSO | 1.064 | 0.156 | | | |




Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Figure S1.** Leave-one-out (LOO) forest plot of univariable Mendelian randomization analysis between the abundance of nine cathepsins (cathepsin B, E, F, G, H, L2, O, S, and Z) and the risk of having chronic obstructive pulmonary disease.

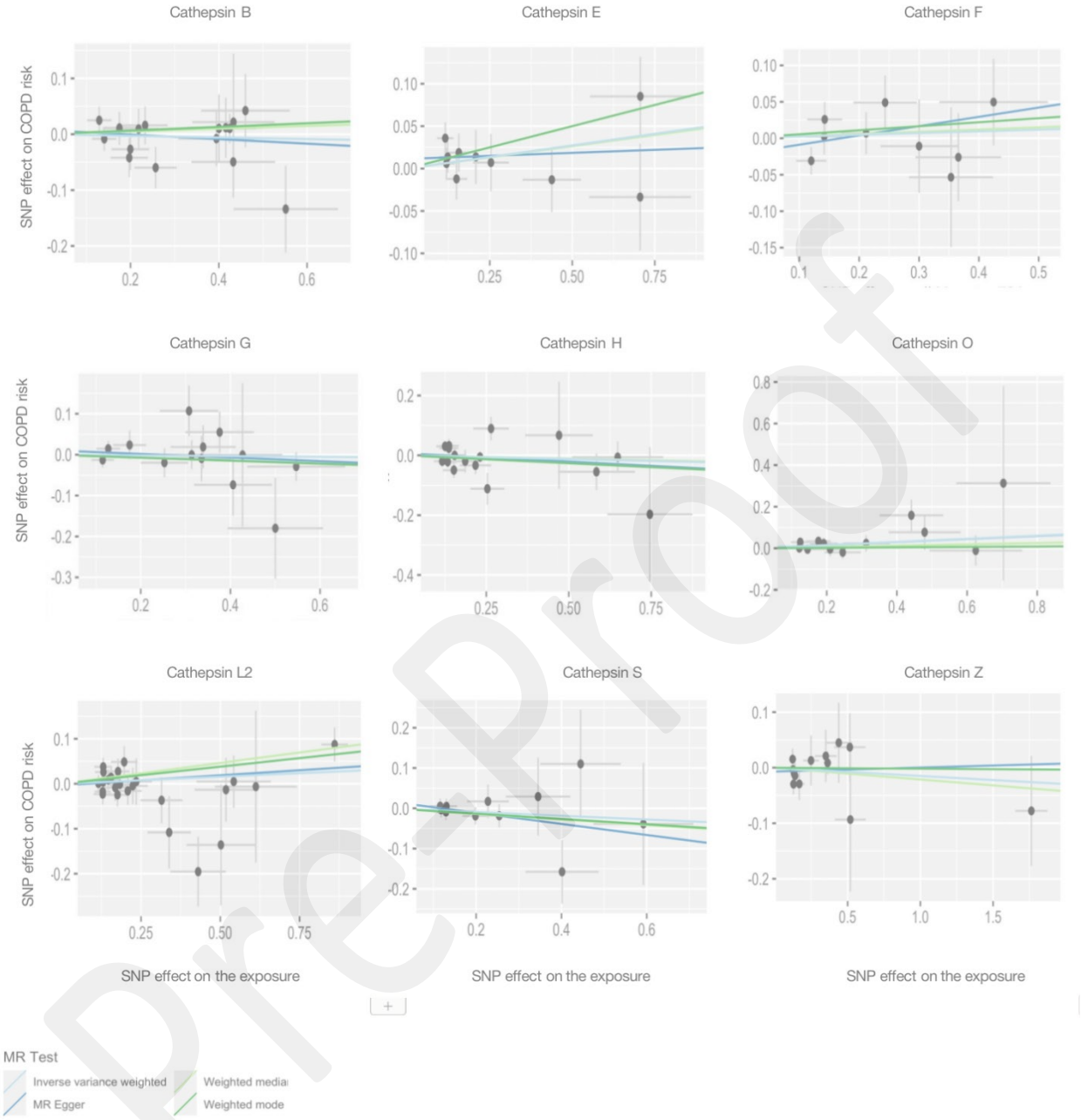






Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626

**Figure S2.** Scatterplots to compare four Mendelian randomization analysis methods.







Published online August 6, 2025 https://doi.org/10.15326/jcopdf.2025.0626